

## LISTING OF THE CLAIMS

Claims 2-6, 8, 10, 12-16, 18, 20, 22-26, 28 and 30 are pending. Please amend claims 2-6, 8, 10, 12-16, 18, 20, 22-26, 28 and 30. No claims are added, canceled, or withdrawn.

The following listing of claims replaces all prior versions and listings of claims in the application.

1. (Canceled)

2. (Currently amended) The A method as recited in of claim 5, wherein the queries comprise a well formed natural language question, a keyword, or a phrase.

3. (Currently amended) The A method as recited in of claim 5, wherein the query cluster is used to disambiguate a word or phrase in a query of the queries.

4. (Currently amended) The A method as recited in of claim 5, further comprising determining that the queries are similar based on similar keyword or phrase composition.

5. (Currently amended) A computer-implemented method for clustering queries, the method comprising:

identifying a same document and/or ~~a plurality of~~ similar documents selected by a user in response to ~~a plurality of~~ queries, the similar documents

1 being determined by evaluating a set of selected similar documents chosen  
2 responsive to queries  $p$  and  $q$  of the queries, wherein documents  $D\_C(.)$  is a subset  
3 of a result list  $D(.)$  according to the following:

$$4 \quad D\_C(p) = \{ d_{p1}, d_{p2}, \dots, d_{pi} \} \subseteq D(p)$$

$$5 \quad D\_C(q) = \{ d_{q1}, d_{q2}, \dots, d_{qj} \} \subseteq D(q);$$

6 wherein similarity based on selection of documents is based on:

7 If  $D\_C(p) \cap D\_C(q) = \{ d_{pq1}, d_{pq2}, \dots, d_{pqk} \} \neq \emptyset$ , then documents  
8  $d_{pq1}, d_{pq2}, \dots, d_{pqk}$  represent a set of common topics of queries  $p$  and  $q$ , and,

9 whereby the similar documents between queries  $p$  and  $q$  is determined by  
10  $D\_C(p) \cap D\_C(q)$ ; and

11 responsive to identifying the same document and/or the similar documents,  
12 generating a query cluster to indicate ~~that the~~ queries  $p$  and  $q$  are similar  
13 independent of whether individual ones of the queries  $p$  and  $q$  comprise similar  
14 composition with respect to other ones of the queries.

15  
16 6. (Currently amended) ~~The~~ A method ~~as recited in~~ of claim 5, further  
17 comprising constructing a thesaurus comprising ~~a plurality of~~ synsets, wherein  
18 each synset comprises one or more query clusters.

19  
20 7. (Canceled)

21  
22 8. (Currently amended) A computer-implemented method for clustering  
23 queries, the method comprising:

24 identifying a same document and/or ~~a plurality of~~ similar documents  
25 selected by a user in response to ~~a plurality of~~ queries, by

determining the similar documents based on a proportionality of commonly selected individual documents as follows, ~~such that~~:

$$similarity_{single\_doc}(p, q) = \frac{RD(p, q)}{Max(rd(p), rd(q))},$$

wherein  $rd(.)$  is the number of clicked documents for a query of the queries, and wherein  $RD(p, q)$  is the number of document selections in common; and

responsive to identifying the same document and/or the similar documents, generating a query cluster to ~~indicate that the queries are similar~~ identify similar queries of the queries independent of whether individual ones of the similar queries comprise ~~similar~~ analogous composition with respect to other ones of the similar queries.

## 9. (Canceled)

10. (Currently amended) A computer-implemented method for clustering queries, the method comprising:

identifying a same document and/or a ~~plurality of~~ similar documents selected by a user in response to a ~~plurality of~~ queries, the similar documents being based on a hierarchical positioning between individual ones of a ~~plurality of~~ documents commonly selected across the queries, wherein  $F(d_i, d_j)$  is a lowest common parent node for documents  $d_i$  and  $d_j$ , wherein  $L(x)$  is a level of a node  $x$ , wherein  $L\_Total$  identifies a total number of levels in a hierarchy, and wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that wherein}$$

$s(d_i, d_j) = 1$ ; and  $s(d_i, d_j) = 0$  if  $F(d_i, d_j) = \text{root}$ ; incorporating  $s(d_i, d_j)$  into a calculation of query similarity, wherein  $d_i$  ( $1 \leq i \leq m$ ) and  $d_j$  ( $1 \leq j \leq n$ ) be a set of selected documents for queries  $p$  and  $q$  respectively as follows ~~such that~~:

$$\text{similarity}_{\text{hierarchy}}(p, q) = \frac{1}{2} \times \left( \frac{\sum_{i=1}^m (\max_{j=1}^n s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^n (\max_{i=1}^m s(d_i, d_j))}{rd(q)} \right); \text{ and}$$

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that ~~the~~ queries p and q are similar independent of whether individual ones of the queries p and q comprise similar composition with respect to other ones of the queries.

11. (Canceled)

1           12. (Currently amended) ~~The Computer-readable media~~ computer-  
2 readable medium ~~as recited in~~ of claim 15, wherein the queries comprise a well  
3 formed natural language question, a keyword, or a phrase.

4  
5           13. (Currently amended) ~~The Computer-readable media~~ computer-  
6 readable medium ~~as recited in~~ of claim 15, wherein the query cluster is used to  
7 disambiguate a word or phrase in a query of the queries.

8  
9           14. (Currently amended) ~~The Computer-readable media~~ computer-  
10 readable medium ~~as recited in~~ of claim 15, wherein the ~~computer-executable~~  
11 computer-program instructions further comprise instructions for determining that  
12 the queries are similar based on similar keyword or phrase composition.

13  
14           15. (Currently amended) ~~Computer-readable media~~ A computer-  
15 readable medium comprising ~~computer-executable~~ computer-program instructions  
16 executable by a processor for identifying similar queries, the ~~computer-executable~~  
17 computer-program instructions comprising instructions for:

18           identifying a same document and/or ~~a plurality of~~ similar documents  
19 selected by a user in response to ~~a plurality of~~ queries, the similar documents  
20 being determined by evaluating a set of selected similar documents chosen  
21 responsive to queries  $p$  and  $q$  of the queries, wherein documents  $D\_C(.)$  is a subset  
22 of a result list  $D(.)$  according to the following:

$$D\_C(p) = \{ d_{p1}, d_{p2}, \dots, d_{pi} \} \subseteq D(p)$$

$$D\_C(q) = \{ d_{q1}, d_{q2}, \dots, d_{qj} \} \subseteq D(q);$$

wherein similarity based on selection of documents is based on:

If  $D\_C(p) \cap D\_C(q) = \{ d_{pq1}, d_{pq2}, \dots, d_{pqk} \} \neq \emptyset$ , then documents  $d_{pq1}, d_{pq2}, \dots, d_{pqk}$  represent a set of common topics of queries  $p$  and  $q$ , and,

whereby the similar documents between queries  $p$  and  $q$  is determined by  $D\_C(p) \cap D\_C(q)$ ; and

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that the queries are similar independent of whether individual ones of the queries comprise similar composition with respect to other ones of the queries.

16. (Currently amended) ~~The Computer-readable media computer-readable medium as recited in~~ of claim 15, wherein the ~~computer-executable computer-program~~ instructions further comprise instructions for constructing a thesaurus comprising a plurality of synsets, wherein each synset comprises one or more query clusters.

17. (Canceled)

1        18.    (Currently amended)    ~~Computer-readable media~~    A computer-  
2    readable medium comprising ~~computer-executable~~ computer-program instructions  
3    executable by a processor for identifying similar queries, the ~~computer-executable~~  
4    computer-program instructions comprising instructions for:

5        identifying a same document and/or ~~a plurality of~~ similar documents  
6    selected by a user in response to ~~a plurality of~~ queries, the similar documents  
7    being determined based on a proportionality of commonly selected individual  
8    documents as follows, ~~such that~~:

$$\text{similarity}_{\text{single\_doc}}(p, q) = \frac{RD(p, q)}{\text{Max}(rd(p), rd(q))},$$

10    wherein  $rd(.)$  is the number of clicked documents for a query of the queries, and  
11    wherein  $RD(p, q)$  is the number of document selections in common; and

12        responsive to identifying the same document and/or the similar documents,  
13    generating a query cluster to indicate ~~that the~~ queries are similar independent of  
14    whether individual ones of the queries comprise similar composition with respect  
15    to other ones of the queries.

17        19.    (Canceled)

20. (Currently amended) ~~Computer-readable media~~ A computer-readable medium comprising ~~computer-executable computer-program~~ instructions executable by a processor for identifying similar queries, the ~~computer-executable computer-program~~ instructions comprising instructions for:

identifying a same document and/or ~~a plurality of~~ similar documents selected by a user in response to ~~a plurality of~~ queries, the similar documents being based on a hierarchical positioning between individual ones of ~~a plurality of~~ documents commonly selected across the queries, wherein  $F(d_i, d_j)$  is a lowest common parent node for documents  $d_i$  and  $d_j$ , wherein  $L(x)$  is a level of a node  $x$ , wherein  $L\_Total$  identifies a total number of levels in a hierarchy, and wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that wherein}$$

$s(d_i, d_j) = 1$ ; and  $s(d_i, d_j) = 0$  if  $F(d_i, d_j) = \text{root}$ ; incorporating  $s(d_i, d_j)$  into a calculation of query similarity, wherein  $d_i$  ( $1 \leq i \leq m$ ) and  $d_j$  ( $1 \leq j \leq n$ ) be a set of selected documents for queries  $p$  and  $q$  respectively as follows ~~such that~~:

$$\text{similarity}_{\text{hierarchy}}(p, q) = \frac{1}{2} \times \left( \frac{\sum_{i=1}^m (\max_{j=1}^n s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^n (\max_{i=1}^m s(d_i, d_j))}{rd(q)} \right); \text{ and}$$

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate ~~that the similar~~ queries ~~are similar~~ independent of whether individual ones of the similar queries comprise similar composition with respect to other ones of the similar queries.



1           21.   (Canceled)

2  
3           22.   (Currently amended) A The computing device ~~as recited in~~ of claim  
4 25, wherein the queries comprise a well formed natural language question, a  
5 keyword, or a phrase.

6  
7           23.   (Currently amended) A The computing device ~~as recited in~~ of claim  
8 25, wherein the query cluster is used to disambiguate a word or phrase in a query  
9 of the queries.

10  
11           24.   (Currently amended) A The computing device ~~as recited in~~ of claim  
12 25, wherein the computer-executable instructions further comprise instructions for  
13 determining that the queries are similar based on similar keyword or phrase  
14 composition.

15  
16           25.   (Currently amended) A computing device comprising:  
17 a processor coupled to a memory, the memory comprising computer  
18 executable instructions, the processor being configured to fetch and execute the  
19 computer-executable instructions for:

20 identifying a same document and/or ~~a plurality of~~ similar documents  
21 selected by a user in response to ~~a plurality of~~ queries, the similar documents  
22 being determined by evaluating a set of selected similar documents chosen  
23 responsive to queries  $p$  and  $q$  of the queries, wherein documents  $D\_C(.)$  is a subset  
24 of a result list  $D(.)$  according to the following:

$$D\_C(p) = \{ d_{p1}, d_{p2}, \dots, d_{pi} \} \subseteq D(p)$$

$$D\_C(q) = \{ d_{q1}, d_{q2}, \dots, d_{qj} \} \subseteq D(q);$$

wherein similarity based on selection of documents is based on:

If  $D\_C(p) \cap D\_C(q) = \{ d_{pq1}, d_{pq2}, \dots, d_{pqk} \} \neq \emptyset$ , then documents

$d_{pq1}, d_{pq2}, \dots, d_{pqk}$  represent a set of common topics of queries  $p$  and  $q$ , and,

whereby the similar documents between queries  $p$  and  $q$  is determined by

$D\_C(p) \cap D\_C(q)$ ; and

responsive to identifying the same document and/or the similar documents,

generating a query cluster to indicate that the queries are similar independent of

whether individual ones of the queries comprise similar composition with respect

to other ones of the queries.

26. (Currently amended) A The computing device ~~as recited in~~ of claim 25, wherein the computer-executable instructions further comprise instructions for constructing a thesaurus comprising ~~a plurality of~~ synsets, wherein each synset comprises one or more query clusters.

27. (Canceled)

28. (Currently amended) A computing device comprising :  
a processor coupled to a memory, the memory comprising computer executable instructions, the processor being configured to fetch and execute the computer-executable instructions for:

identifying a same document and/or ~~a plurality of~~ similar documents selected by a user in response to ~~a plurality of~~ queries, the similar documents

1 being determined based on a proportionality of commonly selected individual  
2 documents as follows, ~~such that~~:

$$3 \quad \text{similarity}_{\text{single\_doc}}(p, q) = \frac{RD(p, q)}{\text{Max}(rd(p), rd(q))},$$

4 wherein  $rd(.)$  is the number of clicked documents for a query of the queries, and  
5 wherein  $RD(p, q)$  is the number of document selections in common; and

6 responsive to identifying the same document and/or the similar  
7 documents, generating a query cluster to indicate ~~that the~~ similar queries ~~are~~  
8 ~~similar~~ independent of whether individual ones of the similar queries comprise  
9 similar composition with respect to other ones of the similar queries.  
10

11 **29.** (Canceled)  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

30. (Currently amended) A computing device ~~comprising~~: comprising:  
a processor coupled to a memory, the memory comprising computer  
executable instructions, the processor being configured to fetch and execute the  
computer-executable instructions for:

identifying a same document and/or a ~~plurality of~~ similar documents  
selected by a user in response to a ~~plurality of~~ queries, the similar documents  
being based on a hierarchical positioning between individual ones of a ~~plurality of~~  
documents commonly selected across the queries, wherein  $F(d_i, d_j)$  is a lowest  
common parent node for documents  $d_i$  and  $d_j$ , wherein  $L(x)$  is a level of a node  $x$ ,  
wherein  $L\_Total$  identifies a total number of levels in a hierarchy, and wherein a  
similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that wherein}$$

$s(d_i, d_i) = 1$ ; and  $s(d_i, d_j) = 0$  if  $F(d_i, d_j) = \text{root}$ ; incorporating  $s(d_i, d_j)$   
into a calculation of query similarity, wherein. (Original)  $d_i$  ( $1 \leq i \leq m$ ) and  $d_j$  ( $1 \leq j \leq$   
 $n$ ) be a set of selected documents for queries  $p$  and  $q$  respectively as follows ~~such~~  
~~that~~:

$$\text{similarity}_{\text{hierarchy}}(p, q) = \frac{1}{2} \times \left( \frac{\sum_{i=1}^m (\max_{j=1}^n s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^n (\max_{i=1}^m s(d_i, d_j))}{rd(q)} \right); \text{ and}$$

responsive to identifying the same document and/or the similar  
documents, generating a query cluster to indicate ~~that the~~ similar queries are  
~~similar~~ independent of whether individual ones of the similar queries comprise  
similar composition with respect to other ones of the similar queries.